

Figure 1: cDNA sequence (and encoded amino acid sequence) of Cynomologous monkey Cathepsin S

```

1      AGTTGAAC TGAAATCTCCCTGCTGCTGCTTGAAATCTTAGAAGAGAGGCCATCAATTCA
61     AGGATTCTTACTGTAGGAGCACCTGCTGGTCTATCACATGAAACAGCTGGTTGTG
                           M K Q L V C V
◆
121     CTCTGGTGTGCTCCTCTGCAGTGGCGCAGTTGCATAAAAGATCCTACCCCTGGATCATCAC
      L L V C S S A V A Q L H K D P T L D H H
181     TGGCATCTCTGGAAAGAAAACCTATGGCAAACAAATACAAGGAAAAGAATGAAGAACAGTA
      W H L W K K T Y G K Q Y K E K N E E A V
241     CGACGTCTCATCTGGGAAAAGAATCTAAAGTTGTGATGCTCATACCTGGAGCATCA
      R R L I W E K N L K F V M L H N L E H S
301     ATGGGAATGCACTCATATGATCTGGGATGAACACCTGGGAGACATGACCAGTGAAGAA
      M G M H S Y D L G M N H L G D M T S E E
361     GTGATGCTTTGAGAGTCCCTGAGAGTCCAGGAGTCAGAGAAATATCACATAT
      V M S L M S S L R V P S Q W Q R N I T Y
↓
421     AAGTCAAACGCTAATCAGATATTGCCGGATTCTGTGGACTGGAGAGAGAAAGGGTGTGTT
      K S N A N Q I L P D S V D W R E K G C V
481     ACTGAAGTGAAAATATCAAGGTTCTTGTGGCTTGGCTTCAGTGCTGTGGGCC
      T E V K Y Q G S C G A C W A F S A V G A
541     CTGGAAGCACAGCTGAAGCTGAAAACAGGAAGCTGGTGTCTCTCAGTGCCAGAACCTG
      L E A Q L K L T G K L V S L S A Q N L
601     GTGGATTGCTCAACTGAAAATATGGAAACAAAGGCTGCAATGGCTTCATGACAAGG
      V D C S T E K Y G N K G C N G G F M T R
661     GCTTTCCAGTACATCATTGATAACAACGGCATCGACTCAGACGCTTCCATCCCTACAAA
      A F Q Y I I D N N G I D S D A S Y P Y K
721     GCCACGGATCAGAAGTGTCAATATGACTCAAATATCGTGCACATGTTCAAAGTAT
      A T D Q K C Q Y D S K Y R A A T C S K Y
781     ACTGAACCTCCATTGAGAGAAGATGTCTGAAAGAAGTTGTGGCAATAAAGGCCA
      T E L P Y G R E D V L K E V V A N K G P
841     GTGTCCTGTTGGTGTGGATGGAGCTCATCCTCTTCTCCCTACAGAAGTGGTGTCTAC
      V S V G V D A S H P S F F L Y R S G V Y
901     TATGAACCATTCTGTACTCAGAATGTGAATCATGGTGTACTTGTGGTGGCTATGGTGT
      Y E P S C T Q N V N H G V L V V G Y G V
961     CTTAACGGGAAAGAATACTGGCTTGTGAAAACAGCTGGGCCAACCTTGGTGAAGAA
      L N G K E Y W L V K N S W G R N F G E E
1021    GGATATATCGGATGGCAAGAATAAAGGAATCATTGTGGGATTGCTAGTTCCCTCT
      G Y I R M A R N K G N H C G I A S F P S
1081    TACCCAGAAAATCTAGAGAGGATCTTCTTTATAACAAATCAAGAAAATATGAAGCAC
      Y P E I *
1141    TTTCTCTTAACCTAATTTTCTGCTGTATCCAGAAGAAATAATTGTGTACGATTAATG
1201    TGTATTTACTGTACTAATTTAAAAATATAGTTTGAGGCCGGCACGGTGGCTACGCCCTG
1261    TAATCCCGACTTGGGAGGGCAAGGCAGGCATATCAACCTGAGGCCAGGAGTTAAAGA
1321    CAGGCCCTGGCTAACATGGTCAAACCCCGTCTACTAAAAAATACAAAACATTAGCCGAGC
1381    GTAATGGTGCATGCCGTAAATCCCAGCTACTGGGAGGCTGAAGCACAAGATTCTGAA
1441    CCCAAGGGTTGAGGCTGTGGTGAAGCTGAGACCCACACTGTACTCCAGCCTGGACAC
1501    AGAGTGGAGACTCTGTTCAAAAAAACAGAAAAGATAATATAGTTGATTCTTCGTTTT
1561    TAAAATTTGCAACCTCAGGAAAAGTTGCTAAGTAAATTAGTAGTGTACTATAGATAT
1621    AACTGTATAAAAATGTTCAACCTAAAACATCTGTCAATTGTTCAATTGTTTATTTATA
1681    CTCTTGTCTTTTAAGACCCCTGATAGCCTTTGTAACTTGTGATGGCTAAAAGTACTTA
1741    ATAAATCTGCCATTCAAATTCAAAAAAAAAAAAAAAAAAAAAAA

```

SEQ ID NO:2

SEQ ID NO:1

◆ = cleavage site for pre-sequence

↓ = cleavage site for pro-sequence

* = stop codon

FIGURE 1

Figure 2: Alignment of cathepsin S amino acid sequences from Cynomologous monkey (SEQ ID NO:2, as shown in Fig. 1), human (Genbank accession number BC002642.2), monkey sml cells (*Saimiri boliviensis*), dog (*Canis familiaris*; Genbank accession number AY156692), and mouse (*Mus musculus*; Genbank gi3850787), followed by a consensus sequence.

“cyno cats preprotein” = Cynomologous monkey = SEQ ID NO:2

“smlmonkey cats” = monkey sml cells (*Saimiri boliviensis*) = SEQ ID NO:3

“hcats preproto.txt” = human (*Homo sapiens*) = SEQ ID NO:4

“dog cats” = dog (*Canis familiaris*) = SEQ ID NO:5

“mouse cats prepro” = mouse (*Mus musculus*) = SEQ ID NO:6

(consensus sequence = SEQ ID NO:7)

FIGURE 2A

ClustalW Formatted Alignments

<i>cyno cats preprotein</i> <i>smlmonkey cats</i> <i>hcats preproto.txt</i> <i>dog cats</i> <i>mouse catsprepro</i>	M K Q L V C V L V C S S A V A Q L H K D P T L D H H W H L W K K T Y G K Q Y K M K Q L V C V L F V C S S A V T Q L H K D P T L D H H W N L W K K T Y G K Q Y K M K R L V C V L L V C S S A V A Q L H K D P T L D H H W H L W K K T Y G K Q Y K M K W L V G L L L P L C S Y A V A Q V H K D P T L D H H W N L W K K T Y S K Q Y K M A V L D A P G V L C G N G A T A E R - - P T L D H H W D L W K K T H E K E Y K M . L V C V L . V C S S A V A Q L H K D P T L D H H W . L W K K T Y G K Q Y K																																							
	10										20										30										40									
	50										60										70										80									
	90										100										110										120									
	130										140										150										160									
170										180										190										200										
210										220										230										240										
250										260										270										280										
290										300										310										320										
330										340										350										360										

FIGURE 2B

Figure 3: Purification of monkey Cathepsin S (Gelcode stained gel)

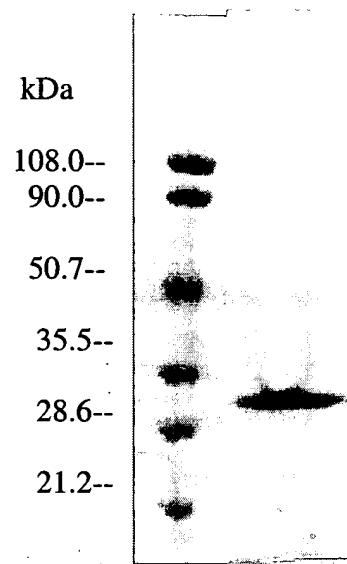


FIGURE 3

BEST AVAILABLE COPY